

10/630,992 PP-89

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 April 2001 (26.04.2001)

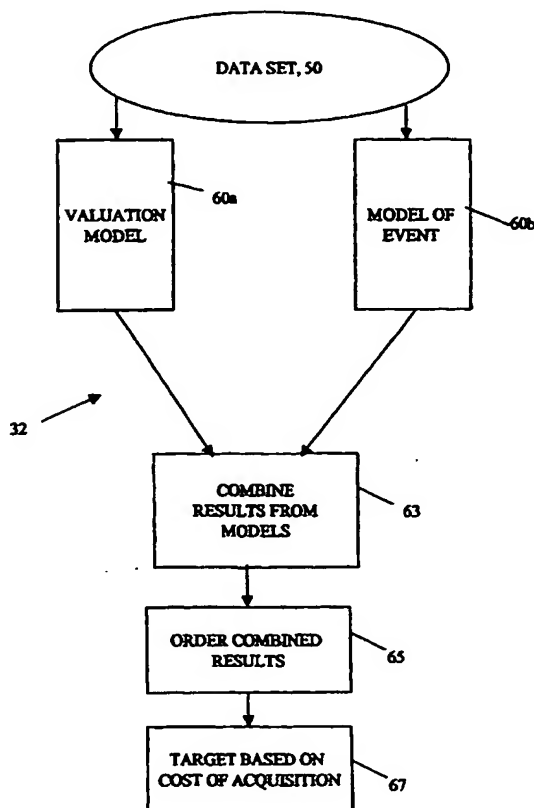
PCT

(10) International Publication Number
WO 01/29692 A2

- (51) International Patent Classification⁷: G06F 17/00 (71) Applicant (for all designated States except US): UNICA TECHNOLOGIES, INC. [US/US]; 55 Old Bedford Road, Lincoln, MA 01773 (US).
- (21) International Application Number: PCT/US00/41232
- (22) International Filing Date: 18 October 2000 (18.10.2000) (72) Inventors; and (75) Inventors/Applicants (for US only): LEE, Yuchun [US/US]; Suite 614, 197 8th Street, Charlestown, MA 02129 (US). CRITES, Robert [US/US]; 1472 Woodhaven Drive, Hummelstown, PA 17036 (US).
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/422,690 21 October 1999 (21.10.1999) US (74) Agent: MALONEY, Denis, G.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: US 09/422,690 (CON) Filed on 21 October 1999 (21.10.1999) (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,

[Continued on next page]

(54) Title: COMBINING VALUE AND PROBABILITY MODELS IN DATABASE MINING



(57) Abstract: A technique to produce a marketing campaign is described. The technique scores a data set of prospects using a plurality of models that estimate state transition probabilities for the prospects, with the models based on samples of potential contacts and their responses, and scores the data set with a plurality of valuation models to determine rewards gained from the prospects in the data set. The model combines the probability of the event occurring for the prospects and values of the prospects to provide targetability value estimates for the prospects by using Markov decision processes from outputs of the scoring and solves the Markov decision processes for the prospects.

WO 01/29692 A2



NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

Published:

*Without international search report and to be republished
upon receipt of that report.*

(84) **Designated States (regional):** ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW). Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM). European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE). OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

COMBINING VALUE AND PROBABILITY MODELS IN DATABASE MINING

BACKGROUND

This invention relates generally to data mining
5 software.

Data mining software extracts knowledge that
may be suggested by a set of data. For example, data
mining software can be used to maximize a return on
investment in collecting marketing data, as well as other
10 applications such as credit risk assessment, fraud
detection, process control, medical diagnoses and so
forth. Typically, data mining software uses one or a
plurality of different types of modeling algorithms in
combination with a set of test data to determine what
15 types of characteristics are most useful in achieving a
desired response rate, behavioral response or other
output from a targeted group of individuals represented
by the data. Generally, data mining software executes
complex data modeling algorithms such as linear
20 regression, logistic regression, back propagation neural
network, Classification and Regression Trees (CART) and
Chi squared Automatic Interaction Detection (CHAD)
decision trees, as well as other types of algorithms on a
set of data.

25

SUMMARY

According to an aspect of the present
invention, a method executed on a computer system
includes scoring prospects or customers with a Markov
30 decision process that models a marketing campaign or
customer relationship management scenario. Aspects of
the invention have the Markov decision process include a

structure for each prospect or customer. The Markov decision process structure includes a set of states, a set of actions, and a determination of which state transition probabilities and rewards vary among the
5 prospects and customers and thus need to be modeled.

According to a further aspect of the present invention, a method executed on a computer system includes scoring a data set of prospects using a plurality of models that estimate state transition
10 probabilities for the prospects, with the models based on samples of potential contacts and their responses, scoring the data set with a plurality of valuation models to determine rewards gained from the prospects in the data set and combining the probability of the event
15 occurring for the prospects and values of the prospects to provide targetability value estimates for the prospects. Aspects have combining use outputs from scoring to produce Markov decision processes and solve the Markov decision processes for the prospects.

20 According to a further aspect of the present invention, a computer program product residing on a computer readable media for developing data for a marketing campaign includes instructions to cause a computer to score a data set of prospects using a
25 plurality of models that estimate state transition probabilities for the prospects, with the models based on samples of potential contacts and their responses and score the data set with a plurality of valuation models to determine rewards gained from the prospects in the
30 data set. The computer program also has instructions to combine the probability of the event occurring for the prospects and values of the prospects to provide

targetability value estimates for the prospects using Markov decision processes.

According to a still further aspect of the invention, a computer program product residing on a computer readable media for developing data for a marketing campaign includes instructions for causing a computer to score prospects or customers with a Markov decision process that models a marketing campaign or customer relationship management scenario.

According to a still further aspect of the invention, a method executed on a computer system includes scoring a data set of prospects using a model that forms a probability of an event occurring for the prospects, with the model based on a sample of potential contacts and their responses and scoring the data set with a valuation model to determine values of the prospects in the data set. The method also includes combining the probability of the event occurring for the prospects and values of the prospects to provide targetability value estimates for the prospects and subtracting from the targetability value estimates costs associated with acquiring the prospects as customers, to provide final targetability values.

One of more of the following advantages may be provided by one or more aspects of the invention.

The application of probability and value models can be based on Markov decision processes, where probability and value models are used to customize a fixed Markov decision process structure. Many types of marketing decisions about which prospects to target for loss-leader or other acquisition campaigns, which customers to target with retention campaigns, and so

forth, can be formalized as Markov decision processes (MDP's). MDP's have a set of states S , a set of actions A , a state transition probability matrix P , and a reward matrix R . Each prospect or customer within a particular campaign has an associated MDP, with the only difference being the state transition probabilities P and possibly the reward matrix R . Solving the MDP for each customer yields the optimal marketing strategy for that customer.

The data mining software uses Markov processes to combine estimates of value of a prospect with estimates that the prospect will respond in a particular manner to events enabling a marketing organization to produce a worthwhile marketing campaign.

15

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of a computer system executing data mining software including an algorithm that combines probability and value modeling.

20

FIG. 2 is a block diagram of a data set.

FIG. 2A is a diagram of a record.

FIG. 3 is a block diagram of a training process of the data mining software.

FIG. 4 is a block diagram of data mining software that uses combined valuation and probability modeling.

FIG. 4A is a block diagram of data mining software that uses combined valuation and probability modeling and Markov decision processes.

FIGS. 5 and 6 are diagrams of Markov decision processes using state transition probabilities and reward structures.

FIG. 7 is a flow chart showing a technique for applying Markov decision processes to marketing campaigns.

5 DETAILED DESCRIPTION

Referring now to FIG. 1, a computer system 10 includes a CPU 12, main memory 14 and persistent storage device 16 all coupled via a computer bus 18. The system 10 also includes output devices such as a display 20 and a printer 22, as well as user input devices such as a keyboard 24 and a mouse 26. Not shown in FIG. 1 but necessarily included in a system of FIG. 1 are software drivers and hardware interfaces to couple all the aforementioned elements to the CPU 12.

15 The computer system 10 also includes data mining software 30 that includes probability and value modeling software 32. The probability and value modeling software 32 determines the desirability and likely response of a prospect. The data mining software 30 may 20 reside on the computer system 10, as shown, or may reside on a server 28 that is coupled to the computer system 10 in a conventional client-server arrangement. The details on how this data mining software is coupled to this computer system 10 are not important to understand the 25 present invention.

Generally, data mining software 30 executes complex data modeling algorithms such as linear regression, logistic regression, back propagation neural network, Classification and Regression Trees (CART) and 30 Chi squared Automatic Interaction Detection (CHAD) decision trees, as well as other types of algorithms that operate on a data set. Also, the data mining software 30

can use any one of these algorithms with different modeling parameters to produce different results. The data mining software 30 can render a visual representation of the results on the display 20 or
5 printer 22 to provide a decision maker with the results. The results that are returned can be based on different algorithm types or different sets of parameters used with the same algorithm. The results can be returned with or without a visual depiction of the results such as the
10 score itself, calculating an RMS value, and so forth. One approach is to render a graph or other visual depiction of the results.

Referring now to FIGS. 2 and 2A, a data set 50 includes a plurality of records 51. The data set 50
15 generally includes a very large number of such records 51. The records 51 (FIG. 2A) can include an identifier field 53a, as well as one or a plurality of fields 53b corresponding to input variable values that are used in the modeling process 30. One of the input variables is a
20 field that has a value of whether a particular event is "true" or "false" for a prospect corresponding to the record 51. The records 51 also include a plurality of result fields 53c where the modeling process (FIG. 4) records scores. The scores are a measure of the expected
25 behavior of a prospect represented by the record. For example, for a record 51a, the result fields 53c include score fields 64a-67a for a corresponding plurality of models that model the likelihood of the prospect responding to an event. Examples of events include
30 customer attrition, response to loss leader, and new customer acquisition. The data mining software 30 or user randomly selects records from data set 50 to produce

a test sample 54. The marketing organization conducts a test marketing campaign 56 and can send randomly selected prospects some marketing literature, items, etc., that is, anything that might elicit a response. The data mining software 30 records 58 in the records 51 whether the event, i.e., response, was true or false for the prospects in the test sample 54. The data mining software 30 partitions 59 the test sample 54 into two groups of records. One group 54a corresponds to records where the event was "true", e.g., a responder, and the other group 54b corresponds to the event being "false", e.g., a non-responder.

Based on the events recorded, an event model is developed. As will be described below, both groups 54a, 54b are used to train the event model. The event model is used to evaluate the probability of an event being true $p(\text{event}=1)$. The group 54a that is true is also used to train a valuation model. The number of prospects in the test sample 54 can be determined based on the size of the data sample.

Alternatively, the event model can be based on data from a previous marketing effort. An alternative is to train a single valuation model with the entire sample and assign a value of zero to records in the group for which the event is false.

Referring now to FIG. 3, group 54a corresponding to true responses is used to produce and train a valuation model 60a. Valuation model 60a is used to predict a valuation of a likely prospect. Desired outputs are based on valuations observed from group 54a. Data from both groups 54a and 54b are used to produce and train an event model 60b. Model 60b is used to predict

the probability of the event occurring for a particular individual. Desired outputs are 1 for individuals from group 54a, and 0 for individuals from group 54b. The results produced by models 60a, 60b are compared 62 to
5 the test marketing data or test history data to evaluate and correct the models by using supervised learning algorithms.

Referring now to FIG. 4, the valuation model 60a and the event model 60b are used to evaluate records
10 in the data set 50. The valuation model 60a provides an estimate of the value, that is, the attractiveness of a particular prospect; whereas, the event model 60b provides a probability of the event occurring for that prospect. That is, the algorithm uses the valuation
15 model 60a to find the desirability of a prospect and the event model 60b to find the likely response of the prospect. The desirability and likely response are combined 63 by the data mining software 30 to provide a targetability type of measurement. Prospects can be
20 ordered 65 by targetability value. As described below, prospects can also be targeted by taking into consideration 67 the cost of acquisition of the prospect. See for example Equation 5.

Desirability corresponds to how attractive,
25 i.e., how well a prospect fits a given profile. Thus the first model estimates the value of the prospect if the prospect were to respond, while the second model predicts the probability of the prospect responding. The targetability of a prospect is the combination 63 of the
30 two models 60a, 60b.

One approach to combining value and likely response to provide a targetability value is to multiply

the two results. In that case, the targetability value is related to the probability (p) of responding times the value (v) if the prospect responds, as given in Equation 1.

$$\text{value} = a(p) \cdot (v)$$

5

Equation 1

where a can be an empirically determined scaling factor and in the general case is equal to 1.

10 Several examples of events that can be modeled include attrition, loss-leader, and predicting revenue for customer acquisition.

 In an attrition model, the first model 60a models the value of the customer if they remain a
15 customer and the second model 60b models the probability of attrition for the customer. For the case of an attrition model, the combined value and probability i.e., targetability value, is given by Equation 2.

$$\text{true value} = [1 - p(\text{attrition})] * \text{value as customer}$$

20

Equation 2

 In a loss leader model, the first model 60a models the value of the customer if they remain a loyal
25 customer and the second model 60b models the probability that a customer who receives the value of a loss leader i.e., a special price or other consideration that can be used to entice additional purchases, will remain a loyal customer. For a loss-leader, the combined value and

probability i.e., targetability value, is given by
Equation 3.

$$\text{true value} = p(\text{loyal}) * \text{value as customer}$$

Equation 3

5

In a revenue predicting model, the first model
60a models the expected revenue if the customer responds
to a marketing campaign and the second model 60b models
the probability that a customer will respond to the
10 marketing campaign. The combined value and probability,
i.e., targetability value, for a revenue predicting model

$$\text{true value} = p(\text{response}) * \text{revenue if responds}$$

is given by Equation 4.

Equation 4

15

The cost of targeting each prospect can be used
as a basis for decisions about which prospects to target.
Prospects should be targeted in accordance with Equation
5

20

$$\text{Prob}(\text{acquisition}) * \text{Value}(\text{acquisition}) - \text{Cost}(\text{acquisition}) > 0 \quad \text{Equation 5}$$

As illustrated by these examples, the event
model can be used by modeling either an event or its
25 complement, whichever is deemed appropriate.

Referring now to FIG. 4A, an embodiment of a
process to combine value and probability modeling is
shown. Test samples are used to build a plurality of
value models 60a₁-60a_n and probability models 60b₁-60b_n (84

in FIG. 7). These models are combined 63'. In a preferred embodiment, the value models $60a_1-60a_n$ and probability models $60b_1-60b_n$ respectively estimate rewards and state transition probabilities for use in a Markov decision process. The number of value models $60a_1-60a_n$ and probability models $60b_1-60b_n$ built depends on the number of states and actions in the Markov decision process.

Each prospect or customer is scored by the value and probability models, resulting in a Markov decision process that is tailored to that individual prospect or customer, as described below. With this arrangement, more complicated marketing campaigns can be built using a series of value and probability models. For example, customer acquisition and attrition can be combined into a larger model. In a standalone customer acquisition model, the possibility of future attrition may either be ignored, or may be included implicitly in the value model by ascribing lower values to those customers that are more likely to leave. A more comprehensive marketing campaign may model the probability of acquisition multiplied by the value of the acquired customer, and explicitly subtract the probability of attrition multiplied by the value lost by attrition. Since the value lost by attrition depends on the time when the attrition occurs, a series of attrition models may be included to deal with a series of future time periods.

As was mentioned in conjunction with Equation 5, the cost of targeting each prospect can be used as a basis for decisions about which prospects to target. In more complicated marketing campaigns, using Markov

decision processes, this cost can also be taken into consideration, as in Equation 6.

$$\begin{aligned}
 & \text{Prob}(\text{acquisition}) * \text{Value}(\text{acquisition}) - \text{Cost}(\text{acquisition}) - \\
 5 \quad & \text{MIN}_i (\text{Prob}(\text{attrition given intervention } i) * \text{Value_Lost}(\text{attrition}) \\
 & + \text{Cost}(\text{intervention } i)) > 0
 \end{aligned}$$

Equation 6

Equation 6 optimizes over a number of different
 10 possible interventions i , including no intervention.
 Interventions are actions by the marketing organization
 such as offers of discounts or credits, or other programs
 designed to encourage customer loyalty. Probabilities
 associated with each intervention can be modeled by
 15 evaluating the results of the intervention on a random
 test sample from the population.

The application of probability and value models
 can be based on Markov decision processes, where
 probability and value models are used to customize a
 20 fixed Markov decision process structure. Many types of
 marketing decisions about which prospects to target for
 loss-leader or other acquisition campaigns, which
 customers to target with retention campaigns, and so
 forth, can be formalized as Markov decision processes
 25 (MDP's). MDP's have a set of states S , a set of actions
 A , a state transition probability matrix P , and a reward
 matrix R . Each prospect or customer within a particular
 campaign has an associated MDP, with the only difference
 being the state transition probabilities P and possibly
 30 the reward matrix R . Solving the MDP for each customer
 yields the optimal marketing strategy for that customer.

The states S summarize the history of the prospect or customer, along with any other relevant information (e.g. demographics) for predicting their behavior. Given the current state of each prospect or customer, the marketer seeks to select its most profitable action from A . The set of actions may include a number of different offers, catalogs, premiums, discounts, promotions, or collateral materials, as well as the method of contact, including telemarketing or direct mail, or which products to target, for example. Another possible action is to take no action. The state transition probabilities P give the likelihood of each possible behavioral outcome (and resulting state), given the marketing action that was selected. The rewards R associated with each state transition combine the expected costs and revenues for that transition. The probabilities and rewards can be estimated by testing and building probability and value models based on the results.

By allowing for the inclusion of absorbing states with terminal rewards, MDP's offer a great deal of flexibility in the quantity and complexity of state transitions to be modeled. An absorbing state is one that has no further outgoing transitions. When a prospect or customer enters an absorbing state, a terminal or final reward is generated.

Referring now to FIG. 5, an example of a simple customer acquisition model in the form of a Markov decision process (MDP) is shown. In this example, the action to be selected is whether or not to contact a prospect. If the decision is not to contact, MDP transitions the prospect to an absorbing "inactive" state,

e.g., "state 2." If the decision is to contact, there are two possible outcomes: if the prospect responds and becomes a customer, that is modeled as a transition to an absorbing "customer" state, and if the prospect does not
5 respond, that is modeled as a transition to an absorbing "non-responder" state. FIG. 5 displays the "states", "actions", and "state transition probabilities and rewards" of the MDP. In FIG. 5, states 2, 3 and 4 are absorbing states with terminal rewards of zero.

10 It is also possible to have a more realistic expanded model, where any or all of these absorbing states may be divided into additional states. For example, non-responders may be solicited additional times, possibly with more attractive offers, which would
15 add several new states (for those who had rejected n offers, $n = 1, 2, \dots$) to the MDP of FIG. 5. In this case there would also be additional decision points determining when to move a non-responsive prospect to the inactive state.

20 The two (2) percent response rate and net present value listed in FIG. 5 will actually vary depending on each individual prospect. Sample testing can be used to build models of response rates and valuations, and each prospect can be inputted into the
25 models. The output of the models is then used to customize the MDP for that prospect. The states and actions in the MDP will remain fixed across all prospects, but the individual state transition probabilities and rewards will be tailored to each
30 prospect.

A benefit of using MDP's for marketing decisions is that they optimize long-term profits. By looking over

a longer time horizon, they can avoid short-sighted decisions that seem profitable initially, but lead to losses in the longer term. Other benefits are possible.

Referring to FIG. 6, a loss-leader campaign is shown. This example illustrates a situation where a very attractive offer may be made to induce a prospect to become a customer, with the expectation that the future value of the customer will outweigh the initial loss. The MDP's show "States", "Actions", and "State Transition Probabilities and Rewards." However, there may be a great deal of attrition among customers acquired in this way, and the marketer may never make up for its initial losses. Systems that model only the first step (customer acquisition) may make short-sighted decisions because they do not include the predicted future loyalty of customers in their decision-making. By contrast, an MDP can take these factors into account and can model any number of steps into the future. In FIG. 6 state 2, 4, 5, and 6 are absorbing states with terminal rewards of zero.

Once the individualized MDP's have been produced, they can be solved using any of a variety of standard dynamic programming techniques, such as value iteration or policy iteration, or by more recent techniques and approximations, including reinforcement learning and neuro-dynamic programming algorithms. Semi-Markov decision processes can be used instead of strict MDP's if a marketer desires to model customer relationships as discrete event dynamic systems. The discussion about MDP's is applicable to this more general treatment.

Referring now to FIG. 7, a technique for applying Markov decision processes to data mining 63' is shown. The process 63' determines 82 the structure of the MDP (i.e. states and actions) to correspond to a
5 desired campaign or customer relationship management scenario. The process 63' performs 84 sample testing (or gathers historical information) necessary to build probability and value models of the state transition probabilities and rewards. The process 63' scores 86
10 prospects or customers through the models, producing individualized MDP's for each prospect or customer. The process 63' solves 88 the MDP for all customers and takes an optimal action for each customer based on their current state.

15

Other Embodiments

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is
20 intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

25 What is claimed is:

CLAIMS

1. A method executed on a computer system comprising:
scoring prospects or customers with a Markov
5 decision process that models a marketing campaign or customer relationship management scenario.
2. The method of claim 1 wherein the Markov
decision process includes a structure for each prospects
10 or customers.
3. The method of claim 2 wherein the Markov
decision process structure includes a set of states, a
set of actions, and a determination of which state
15 transition probabilities and rewards vary among the prospects and customers and thus need to be modeled.
4. The method of claim 1 further comprising:
testing samples of prospects to model the state
20 transition probabilities and rewards.
5. The method of claim 1 further comprising:
using historical information on samples of
prospects to model the state transition probabilities and
25 rewards.
6. The method of claim 1 wherein information about
each prospect or customer is input to probability and
value models and the output probabilities and values of
30 the models are used to produce a MDP for each prospect or customer.

7. The method of claim 6 wherein the MDP's are solved for optimal actions.
- 5 8. The method of claim 1 wherein the process is a semi-Markov decision process.
9. The method of claim 6 wherein the information about each prospect is provided by sample testing or
10 historical information.
10. A method executed on a computer system comprising:
- scoring a data set of prospects using a
15 plurality of models that estimate state transition probabilities for the prospects, with the models based on samples of potential contacts and their responses;
- scoring the data set with a plurality of
valuation models to determine rewards gained from the
20 prospects in the data set;
- combining the probability of the event occurring for the prospects and values of the prospects to provide targetability value estimates for the prospects.
- 25
11. The method of claim 10 wherein combining comprises:
- using outputs from scoring to produce Markov decision processes; and
30 solving the Markov decision processes for the prospects.

12. The method of claim 10 wherein the training data for the model that measures probability is a random test sample.
- 5 13. The method of claim 10 wherein the training data for the model that measures probability are results from previous marketing campaigns.
- 10 14. The method of claim 10 wherein the valuation models measure reward from a prospect as expected profit.
- 15 15. The method of claim 10 wherein the event model models a probability of attrition of a customer.
- 15 16. The method of claim 10 wherein the event model models a probability of loyalty of a customer.
- 20 17. The method of claim 10 wherein the event model models a probability of response of a customer.
- 20 18. The method of claim 10 wherein the training data for the valuation models are portions of the test sample containing prospects in the same state transition.
- 25 19. The method of claim 10 wherein training data for the event models are samples from the test sample containing prospects in the same state transition.
- 30 20. A computer program product residing on a computer readable media for developing data for a marketing campaign comprises instructions for causing a computer to:

- score a data set of prospects using a plurality of models that estimate state transition probabilities for the prospects, with the models based on samples of potential contacts and their responses;
- 5 score the data set with a plurality of valuation models to determine rewards gained from the prospects in the data set;
- combine the probability of the event occurring for the prospects and values of the prospects to provide
- 10 targetability value estimates for the prospects.
21. The computer program product of claim 20 wherein instructions to combine further comprises instructions to cause the computer to:
- 15 produce Markov decision processes from outputs of the score instructions; and
- solve the Markov decision processes for the prospects.
- 20 22. The computer program product of claim 20 wherein training data for the models that measures probability are results from random test samples.
23. The computer program product of claim 20
- 25 wherein the training data for the models that measures probability are results from previous marketing campaigns.
24. The computer program product of claim 20
- 30 wherein the valuation model measures reward of prospects by expected profit.

25. The computer program product of claim 20
wherein the event model models a probability of attrition
of a customer.

5 26. The computer program product of claim 20
wherein the event model models a probability of loyalty
of a customer.

27. The computer program product of claim 20
10 wherein the event model models a probability of response
of a customer.

28. The computer program product of claim 20
wherein the training data for the valuation models are
15 portions of the test sample containing prospects in the
same state transition.

29. The computer program product of claim 20
wherein training data for the event models are samples
20 from the test sample containing prospects in the same
state transition.

30. A computer program product residing on a
computer readable media for developing data for a
25 marketing campaign comprises instructions for causing a
computer to:

 score prospects or customers with a Markov
decision process that models a marketing campaign or
customer relationship management scenario.

30

31. The computer program product of claim 30 wherein the Markov decision process includes a structure for each prospect or customer.

5 32. The computer program product of claim 31 wherein the Markov decision process structure includes a set of states, a set of actions, and a determination of which state transition probabilities and rewards vary among the prospects and customers and thus need to be
10 modeled.

33. The computer program product of claim 30 wherein information about each prospect or customer is input to probability and value models and output state
15 transition probabilities and rewards of the models are used to produce a MDP for each prospect or customer.

34. A computer system comprising:
a processor;
20 a computer storage media storing a computer program product for marketing comprises instructions for causing a computer to:

score a data set of prospects using a plurality of models that estimate state
25 transition probabilities for the prospects, with the models based on samples of potential contacts and their responses;

score the data set with a plurality of valuation models to determine rewards gained
30 from the prospects in the data set;

combine the probability of the event occurring for the prospects and values of the

prospects to provide targetability value estimates for the prospects;

produce Markov decision processes from outputs of the score instructions; and

5 solve the Markov decision processes for the prospects.

35. The system of claim 30 wherein training data for the models that measure probabilities are results
10 from a random test sample.

36. The system of claim 30 wherein the training data for the models that measure probabilities are results from previous marketing campaigns.

15

37. The system of claim 30 wherein the valuation model measures reward from a prospect as expected profit.

38. The system of claim 30 wherein the event models
20 estimate probabilities of attrition of a customer.

39. The system of claim 30 wherein the event models estimate a probability of loyalty of a customer.

25 40 A method executed on a computer system comprising:

scoring a data set of prospects using a model that forms a probability of an event occurring for the prospects, with the model based on a sample of potential
30 contacts and their responses;

scoring the data set with a valuation model to determine values of the prospects in the data set;

combining the probability of the event.
occurring for the prospects and values of the prospects
to provide targetability value estimates for the
prospects; and

- 5 subtracting from the targetability value
estimates costs associated with acquiring the prospects
as customers, to provide final targetability values.

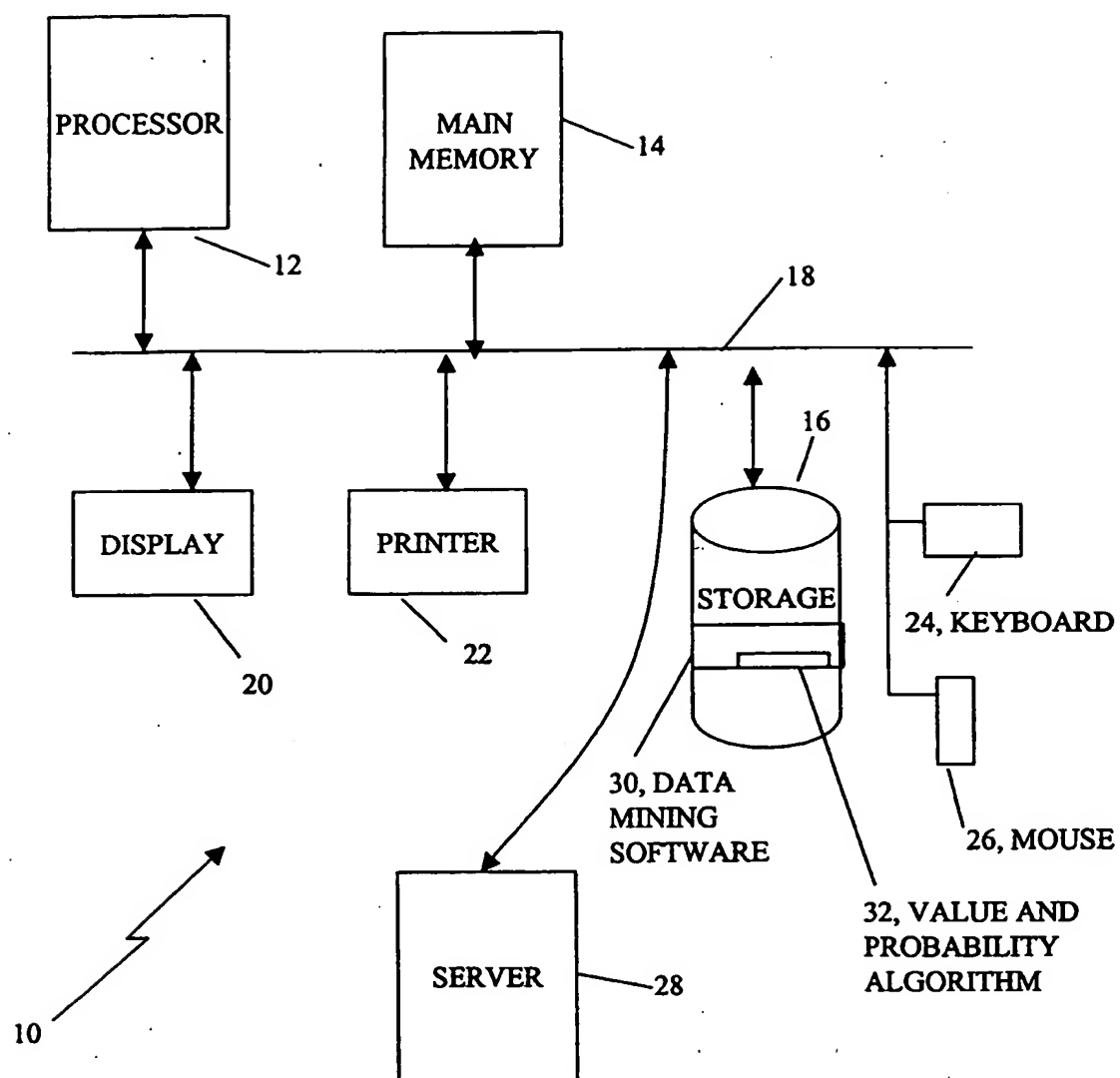
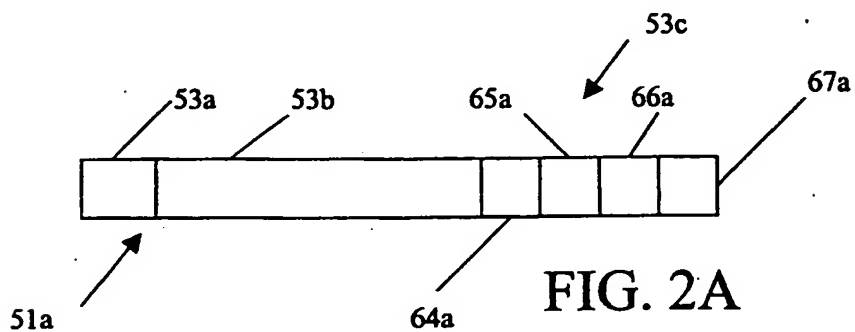
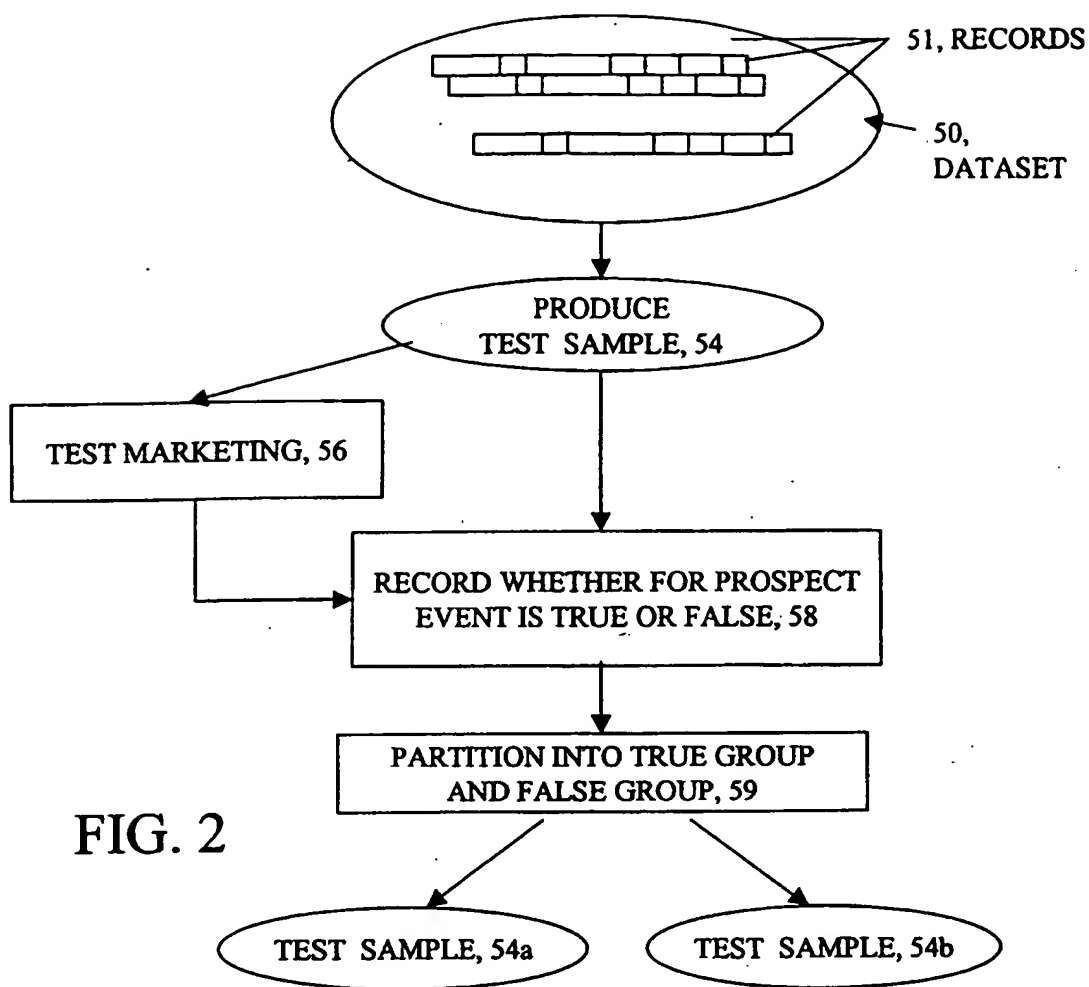


FIG. 1



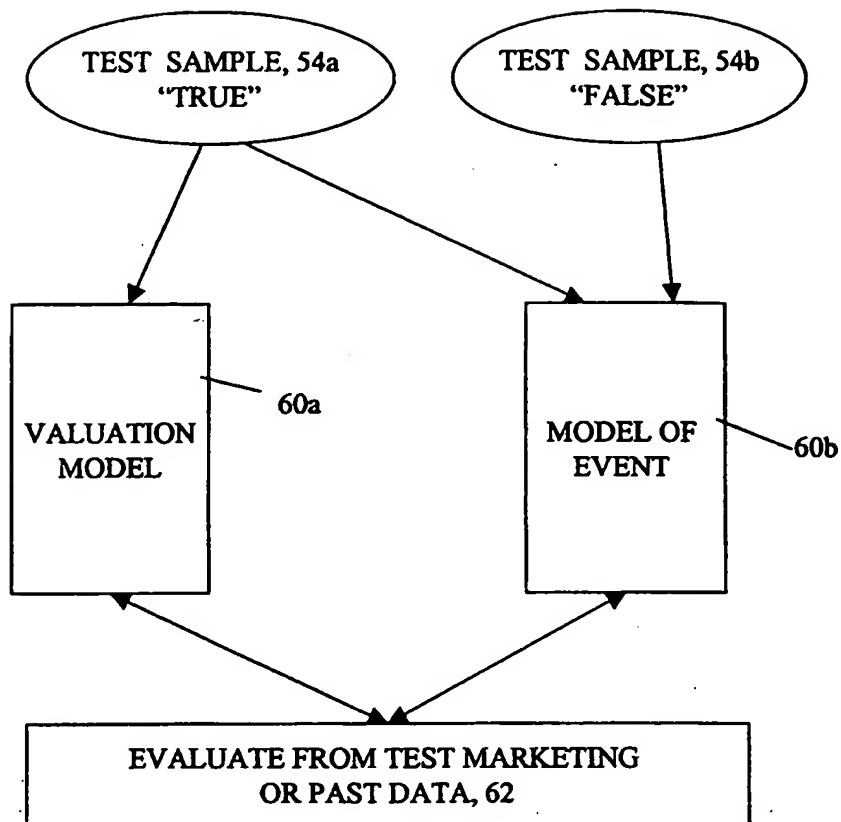
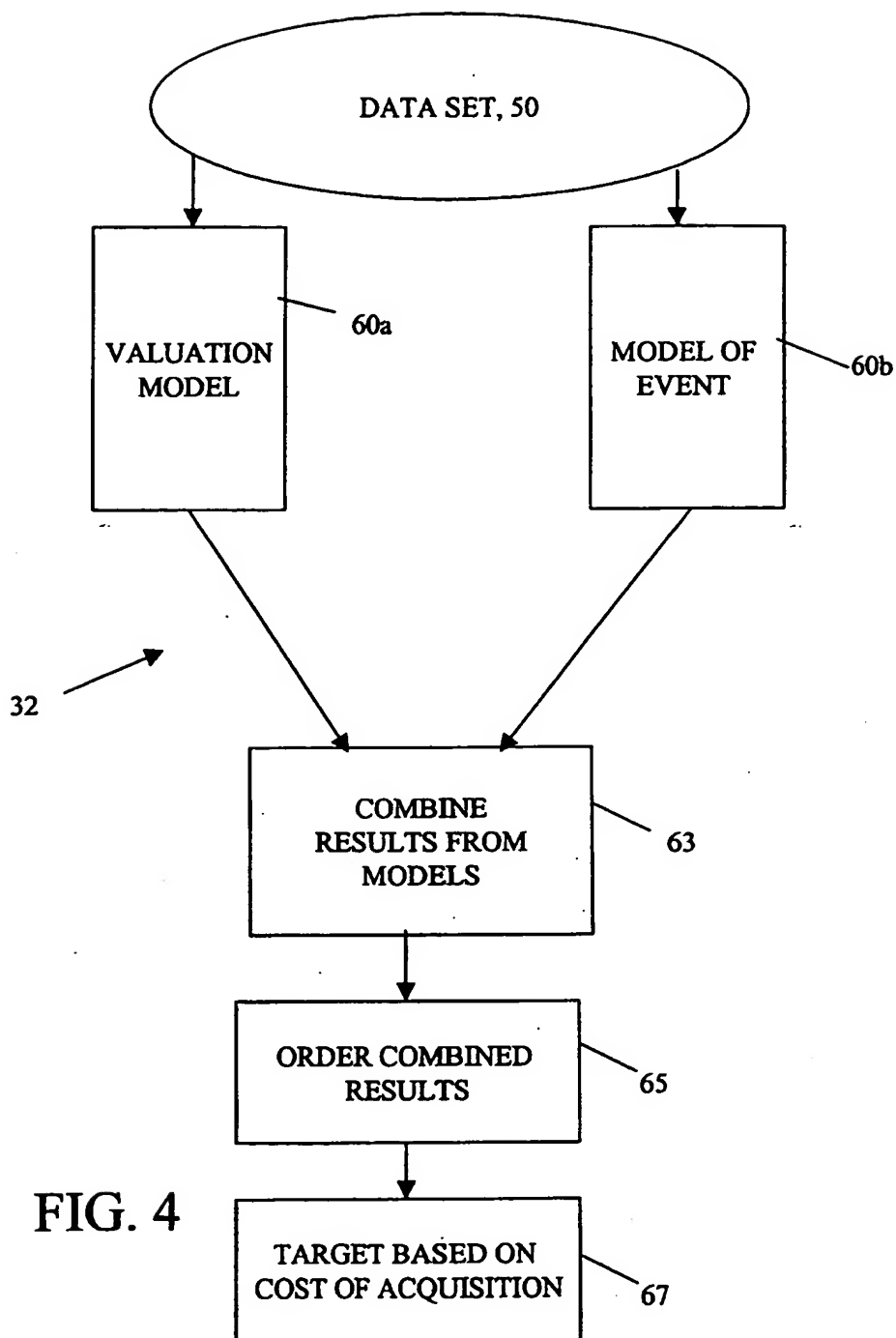


FIG. 3



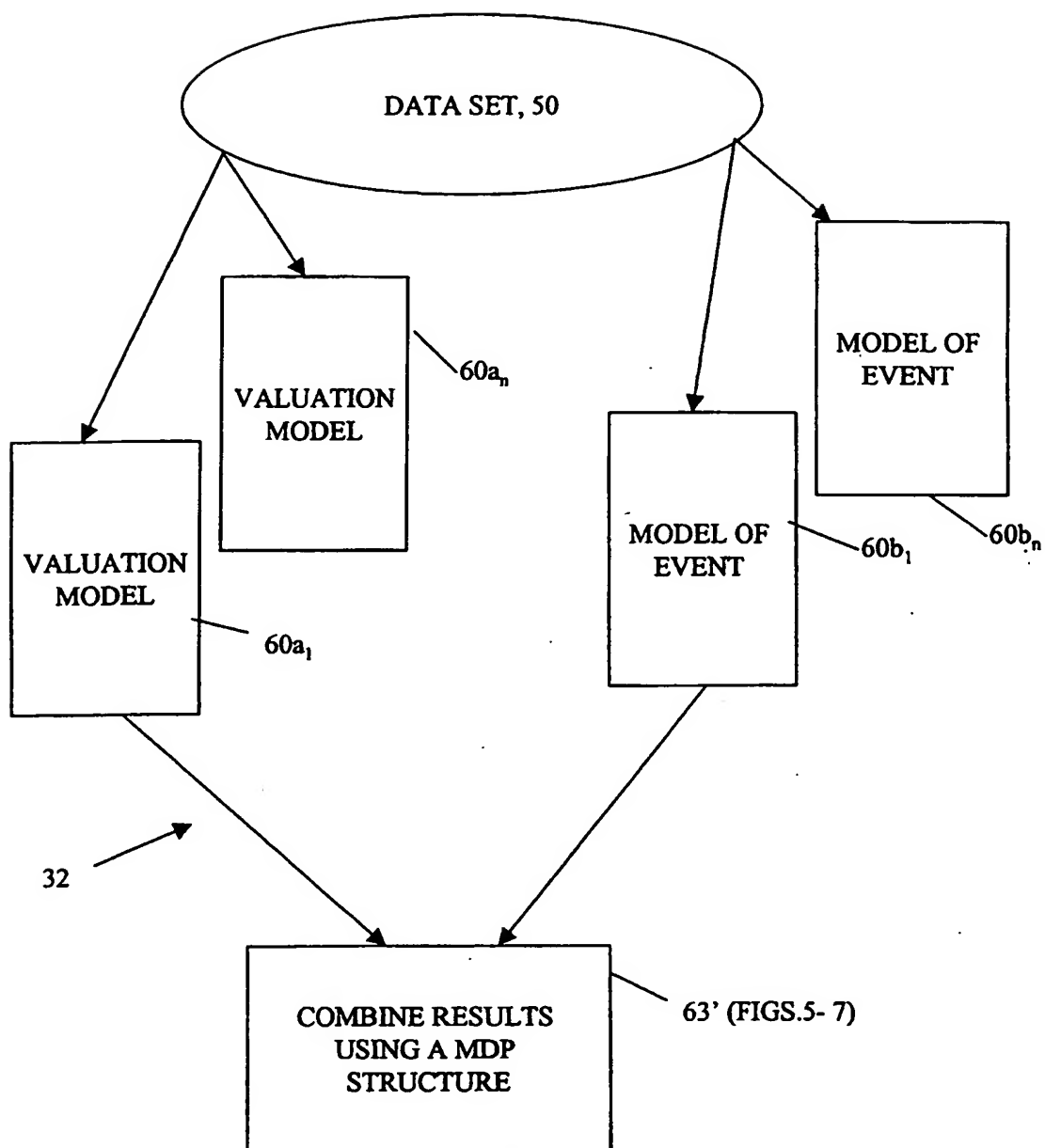
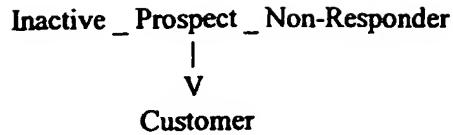


FIG. 4A

FIG. 5

States:

Prospect

Inactive

Customer

Non-Responder

Actions:

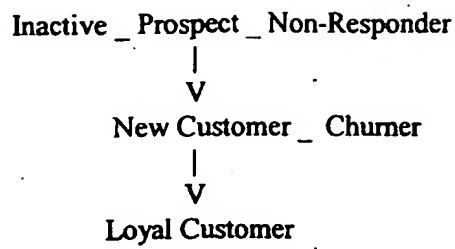
Do not contact

Contact

State Transition Probabilities and Rewards:

Current State	Action	Next State	Probability	Reward
1	1	2	1.00	0.00
1	2	3	.02	net present value – acquisition cost
1	2	4	.98	– acquisition cost

(States 2, 3, and 4 are absorbing states with terminal rewards of zero.)

FIG. 6**States:**

Prospect

Inactive

New Customer

Non-Responder

Churner

Loyal Customer

Actions:

Do not contact

Contact

No action

State Transition Probabilities and Rewards:

Current State	Action	Next State	Probability	Reward
1	1	2	1.00	0.00
1	2	3	.05	– contact cost – loss leader cost
1	2	4	.95	– contact cost
3	3	5	.35	0.00
3	3	6	.65	net present value

(States 2, 4, 5, and 6 are absorbing states with terminal rewards of zero.)

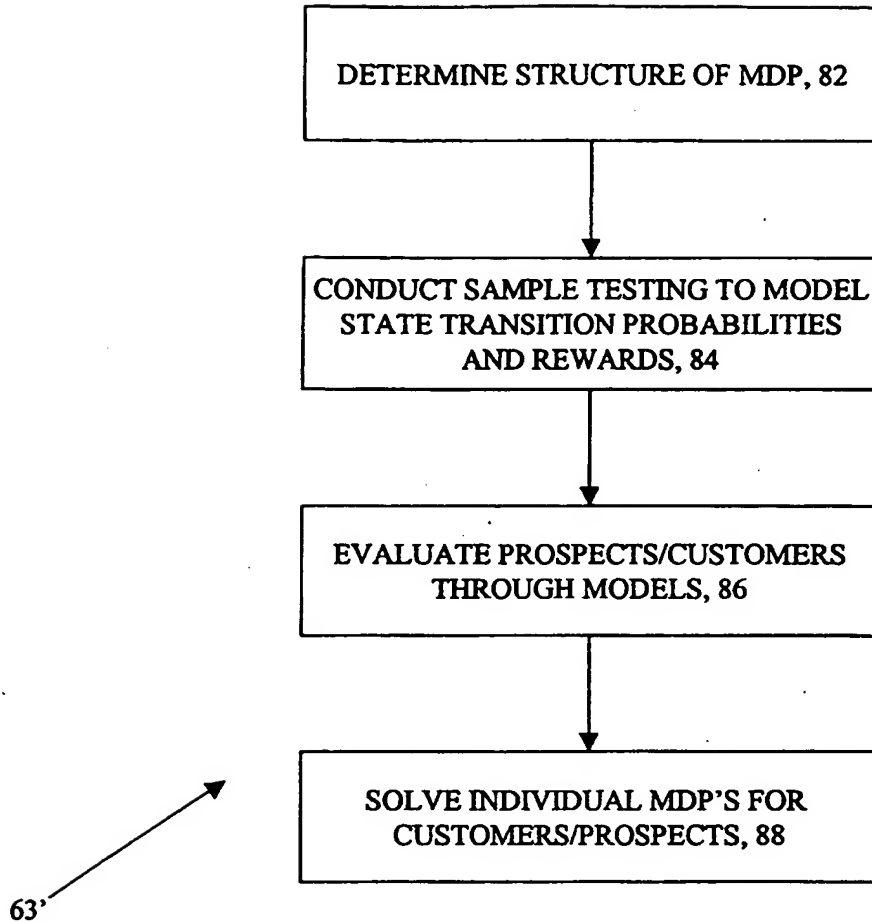


FIG. 7